DOCUMENT RESUME

ED 205 915 CS 006 214

AUTHOR Bruce, Bertram; And Others
TITLE Why Readability Formulas Fail. Reading Education Report No. 28.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY National Inst. of Education (ED), Washington, D.C.
REPORT NO BBN-R-4715
PUB DATE Aug 81
CONTRACT 400-76-0116
NOTE 17p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Cohesion (Written Composition); *Evaluation Methods; Measurement Techniques; Readability; *Readability Formulas; Reading Ability; *Reading Comprehension; Reading Processes; Reading Research; *Validity

ABSTRACT

The failure of readability formulas can be attributed to three weaknesses in the formulas. First, they ignore or violate current knowledge about the reading process. Most formulas affect only sentence length and word difficulty while ignoring factors that influence text comprehensibility, such as cohesion, the number of inferences required, the number of items to remember, complexity of ideas, rhetorical structure, dialect, and required schemata. Nor do they account for reader-specific factors such as interest and the purpose for reading. Second, readability formulas lack solid statistical grounding. The most respected formulas have been validated by test lessons that were intended only as practice exercises, never as measures of text comprehensibility or as indicators of reading ability across age, class, or cultural groups. Third, readability formulas are used inappropriately in two of the contexts in which they appear to be most valuable. Even a formula with some validity, used with appropriate texts and readers, cannot correctly predict how a particular reader will interact with a particular book. (HTH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED205915

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

CENTER FOR THE STUDY OF READING

Reading Education Report No. 28

WHY READABILITY FORMULAS FAIL

Bertram Bruce, Andee Rubin,
and Kathleen Starr

Bolt Beranek and Newman Inc.

August 1981

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

BBN Report No. 4715

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02238

The research reported herein was supported in part by the National Institute of Education under Contract No. HEW-NIE-C-400-76-0116.

CS006214

Why Readability Formulas Fail

Being able to measure the readability of a text with a simple formula is an attractive prospect, and many groups have been using readability formulas in a variety of situations where estimates of text complexity are thought to be necessary. The most obvious and explicit use of readability formulas is by educational publishers designing basal and remedial reading texts; some states, in fact, will consider using a basal series only if it fits certain readability formula criteria. Increasingly, public documents such as insurance policies, tax forms, contracts, and jury instructions must meet criteria stated in terms of readability formulas.

Unfortunately, readability formulas just don't fulfill their promise. This failure can be attributed to three weaknesses in the formulas. From a theoretical point of view, they ignore or violate much of current knowledge about reading and the reading process. Second, their statistical bases are shaky, being at once poorly supported mathematically and difficult to generalize. Finally, as practical tools either for matching children and texts or for providing guidelines for writers they are totally inappropriate. Criticisms such as these have been leveled at readability formulas from many quarters (Gilliland, 1972; Redish, 1979; Kintsch & Vipond, 1977), but the formulas' uses have

expanded in spite of the growing number of papers discussing their weaknesses. We attempt here to categorize and summarize some of the problems with readability formulas and their use.

Factors Not in the Formulas

The first category of problem involves the discrepancy between the characteristics of texts which readability formulas measure and those which we know to influence text comprehensibility. Because most of the formulas include only sentence length and word difficulty as factors, they can account only indirectly for other factors that make a particular text difficult, such as degree of discourse cohesion, number of inferences required, number of items to remember, complexity of ideas, rhetorical structure, dialect, and background knowledge required. Further, because the formulas are measurements based on a text isolated from the context of its use, they cannot reflect such reader-specific factors as motivation, interest, competitiveness, values, and purpose.

Readability formulas fail to account for differences in readers' dialect and cultural backgrounds. For example, a passage in Black Vernacular from the Bridge series (Simpkins, 1977), a cross-cultural reading program, starts:

> Willie went and got hisself a lightweight gig. The gig wasn't saying too much. It wasn't paying nothing but chump change.

Readers familiar with this form of Black Vernacular find the passage relatively simple. Others can infer the meanings of individual words only with difficulty.

Because they view texts so narrowly, readability formulas also fail to measure the effect of the context in which a passage is read. A health information sheet describing the concept and treatment of hypertension, for example, may communicate quite effectively if a patient has enough time to read it and feels comfortable asking a physician for clarification. In a rushed, brusque encounter, however, the document would be much less comprehensible.

Lack of Statistical Basis

Despite the shortcomings of readability formulas on theoretical grounds, strong empirical evidence of their predictive value might justify their use for some tasks. Unfortunately, when such evidence is examined, the second major problem with readability formulas--their lack of solid statistical grounding--becomes apparent. Many of the hundreds of formulas in existence were validated only in terms of earlier

formulas. The early formulas, in turn, were validated using the McCall-Crabbs Standard Test Lessons in Reading (McCall & Crabbs, 1950, 1961). But the McCall-Crabbs lessons were intended only as practice exercises in reading, never as measures of comprehension or text comprehensibility; nor were they intended to be general indicators of reading ability across age, class, or cultural groups. Nevertheless, the most respected formulas have all used the McCall-Crabbs lessons as the criterion of difficulty (Stevens, 1980).

Spache (1978), a readability formula designer, stated the problem succinctly:

> The reading level given by the formula should mean that a child with that level of reading ability could read the book with adequate comprehension and a reasonable number of oral reading errors. This assumption has seldom if ever been tested in the development of this and other readability formulas (emphasis added).

While validation studies were generally not performed in the course of developing readability formulas, a fair number were done after the fact. In a comprehensive review of such studies, Klare (1976) noted that 39 of 65 studies demonstrated a positive correlation between readability formula estimates of difficulty

and reader performance on independent criteria such as reading speed or comprehension. However, even this unconvincing performance is undercut by his observation that positive results are more likely to be reported in journals than negative ones and by the fact that when comprehension, rather than reading speed, is used as the independent measure of text difficulty, only half of the studies indicated positive correlations with readability formula estimates. Lockman (1957) computed Flesch Reading Ease scores for nine sets of instructions for psychological tests, then had 171 naval cadets rate them on "understandability." The rank-order correlation between the two sets of measurements was -0.65, a strong correlation but in the wrong direction.

Common sense also leads us to wonder how generalizable readability formula estimates are beyond the precise situation in which they were validated. In 1978 Spache (1978) developed a revised version of his 1953 formula, saying,

> If a readability formula is to continue to reflect accurate estimates of the difficulty of today's books, it, too, must change.

That is, a formula validated with one group of students and one type of texts is found to be invalid for the same types of students and texts as conditions change over a 25-year period. The effects on validity of the formula for readers having

different cultural backgrounds or dialects must be considerably greater.

Inappropriate Use

This leads us to the third general failing of the readability formulas: Their use is inappropriate in two of the contexts in which they seem most valuable. The first of these is the selection of an appropriate text for a child in school. Even if we assume the formulas have some limited validity and even if we are working with appropriate groups of texts and readers, we can never assume that the formula will correctly predict how a particular reader will interact with a particular book.

For example, the book Don't Forget the Bacon (Hutchins, 1976) is a children's book that scores at grade level 2.7 using the Spache (1978) formula. It has mostly one syllable, easy words and short, simple sentences, e.g., "a pile of chairs?". Nevertheless, some children in fourth grade find it difficult to understand because the higher-level structure of the story is complex and subtle. The main character is a small boy given a verbal grocery list by his mother. Understanding the story depends on distinguishing between times the boy is rehearsing the list in order to remember it and times he is repeating the same list in order to figure out what went wrong. Because of this twist, the book may be more complex than its low score implies.

Relying on the formulas either to gauge the book's readability or a child's reading level could be worse than useless.

A second major use for readability formulas has been as guidelines for the simplification of existing texts and documents. Here, too, using these formulas is inappropriate. Although they may, in certain cases, assign reasonable numerical values to texts, they by no means justify modifications of an existing text. Yet, in cases where readability formulas are used, writers naturally tend to write to the formulas. Such prescriptive use magnifies the inaccuracies inherent in the formulas.

Several studies have investigated the effect of using readability formulas to guide text revision. An exercise in rewriting jury instructions demonstrated that the score of revised instructions on a readability measure had little to do with how well they were understood by jurors (Charrow & Charrow, 1979).

A study by Davison, Kantor, Hannah, Hermon, Lutz, and Salzillo (1980) showed that adapting texts in the Science Research Associates Skillbuilders series to fit the formulas was not only ineffective, but, in many cases, actually increased the difficulty of the text. For example, in a passage about trees, the sentence

> If given a chance before another fire comes, the tree will heal its own wounds by growing new bark over the burned part.

was changed to

> If given a chance before another fire comes, the tree will heal its own wounds. It will grow new bark over the burned part.

The modified text contains shorter sentences, so according to most readability formulas it should be easier to read. However, the reader must now make the inference that the new bark is the mechanism by which the tree heals its wounds without an explicit statement of this fact. Thus, the adapted text may actually be more difficult than the original.

Criteria for Applicability

The preceding examples illustrate various ways in which readability formulas yield faulty predictions, or even lead to the writing of passages which are harder to read. As a series of separate examples, they do not show why readability formulas fail nor do they distinguish among different situations in which the formulas might be more or less appropriate. In each case, however, we can point to an assumption about the use of the formulas which has been violated. On the basis of these examples

of readability formula failure, then, we are led to the conclusion that the formulas are valid only if certain conditions hold. Interestingly, similar lists of conditions have been put forth by designers of the formulas themselves. It is becoming increasingly clear that readability formulas should be used only where the following criteria are met:

1. Material may be freely read. Material like captioning for the deaf, which appears on the screen and then disappears after a certain time, cannot be freely read. The time spent on it is limited by external factors, not by choice of the reader.

2. Text is honestly written. The formulas assume that material is not written to satisfy the readability formulas, but rather to satisfy some other communicative goal.

3. Higher-level text structures are irrelevant. The formulas assume that organizational material, information about intentions, goals, etc. need not be specifically taken into account.

4. Purpose in reading is irrelevant. Skimming, test-taking, reading for pleasure, and so on are

all taken to be equivalent in determining the readability of a passage.

5. Statistical averages are meaningful in individual cases. Use of the formulas implies that statistical averages regarding both texts and readers can provide useful information regarding the appropriateness of an individual text for an individual person.

6. Readers of interest are the same as the readers on whom the readability formula was validated. Any attempt to expand the use of the formula to evaluate materials for readers whose background, dialect, purpose in reading, etc. differs from those of the readers used in validation is likely to lead to difficulties.

Unfortunately, it appears that not only some, but nearly all, uses of readability formulas violate the basic assumptions on their applicability. Rigorous adherence to these assumptions effectively prevents use of readability formulas for TV captioning, adaptation, selection of texts for readers of different cultural backgrounds, designing special texts for children, selection of text passages, choosing trade books, or designing remedial readers, and restricts readability formula use

to trivial cases of little import for educational or social policy.

We are left with a question: Are there any areas in which the assumptions about the readability formulas are satisfied and the formulas improve on intuitive estimates of the readability of a text? We think not. The real factors that affect readability are elements such as the background knowledge of the reader relative to the knowledge presumed by the writer, the purpose of the reader relative to the purpose of the writer, and the purpose of the person who is presenting the text to the reader. These factors cannot be captured in a simple formula and ignoring them may do more harm than good.

References

Charrow, R., & Charrow, V. Making legal language understandable: A psycholinguistic study of jury instructions. Columbia Law Review, 1979, 79, 1306-1374.

Davison, A., Kantor, R., Hannah, J., Hermon, G.,Lutz, R., Salzillo, R. Limitations of readability formulas in guiding adaptations of texts (Tech. Rep. No. 162). Urbana: University of Illinois Center for the Study of Reading, March 1980. (ERIC Document Reproduction Service No. ED 184 090)

Gilliland, J. Readability. London: University of London Press Ltd., 1972.

Hutchins, P. Don't forget the bacon. New York: Greenwillow Books, 1976.

Kintsch, W., & Vipond, D. Reading comprehension and readability in educational practice and psychological theory. In Lars-Goran Nilsson (Ed.), Proceedings of the Conference on Memory. Hillsdale, N.J.: Erlbaum, 1977.

Klare, G. R. A second look at the validity of readability formulas. Journal of Reading Behavior, 1976, 8, 129-152.

Lockman, R. F. A note on measuring "understandability." Journal of Applied Psychology, 1957, 40, 195-196.

McCall, W. A., & Crabbs, L. M. Standard test lessons in reading. New York: Teachers College Press, 1950, 1961.

Redish, J. Readability. In D. A. McDonald (Ed.), Drafting documents in plain language. New York: Practicing Law Institute, 1979.

Simpkins, G., Holt, G., & Simpkins, C. Bridge - A cross-culture reading program. Boston: Houghton Mifflin, 1977.

Spache, G. D. Good reading for poor readers (rev. 10th ed.). Champaign, Ill.: Garrard, 1978.

Stevens, K. C. Readability formulae and McCall-Crabbs standard test lessons in reading. The Reading Teacher, January 1980, 413-415.

CENTER FOR THE STUDY OF READING

READING EDUCATION REPORTS

Adams, M. J., Anderson, R. C., & Durkin, D. Beginning Reading: Theory and Practice (No. 3), November 1977. (ERIC Document Reproduction Service No. ED 151 722, 15p., PC-$2.00, MF-$.91)

Adams, M., & Bruce, B. Background Knowledge and Reading Comprehension (No. 13), January 1980. (ERIC Document Reproduction Service No. ED 181 431, 48p., PC-$3.65, MF-$.91)

Anderson, R. C., & Freebody, P. Vocabulary Knowledge and Reading (No. 11), August 1979. (ERIC Document Reproduction Service No. ED 177 470, 52p., PC-$5.30, MF-$.91)

Anderson, T. H. Another Look at the Self-Questioning Study Technique (No. 6), September 1978. (ERIC Document Reproduction Service No. ED 163 441, 19p., PC-$2.00, MF-$.91)

Anderson, T. H., Armbruster, B. B., & Kantor, R. N. How Clearly Written are Children's Textbooks? Or, Of Bladderworts and Alfa (includes a response by M. Kane, Senior Editor, Ginn and Company) (No. 16), August 1980. (ERIC Document Reproduction Service No. ED 192 275, 63p., PC-$5.30, MF-$.91)

Armbruster, B. B., & Anderson, T. H. Content Area Textbooks (No. 23), July 1981.

Asher, S. R. Sex Differences in Reading Achievement (No. 2), October 1977. (ERIC Document Reproduction Service No. ED 146 567, 30p., PC-$3.65, MF-$.91)

Baker, L. Do I Understand or Do I not Understand: That is the Question (No. 10), July 1979. (ERIC Document Reproduction Service No. ED 174 948, 27p., PC-$3.65, MF-$.91)

Bruce, B. What Makes a Good Story? (No. 5), June 1978. (ERIC Document Reproduction Service No. ED 158 222, 16p., PC-$2.00, MF-$.91)

Bruce, B. A New Point of View on Children's Stories (No. 25), July 1981.

Bruce, B., & Rubin, A. Strategies for Controlling Hypothesis Formation in Reading (No. 22), June 1981.

Bruce, B., Rubin, A., & Starr, K. Why Readability Formulas Fail (No. 28), August 1981.

Collins, A., & Haviland, S. E. Children's Reading Problems (No. 8), June 1979. (ERIC Document Reproduction Service No. ED 172 188, 19p., PC-$2.00, MF-$.91)

Davison, A. Readability--Appraising Text Difficulty (No. 24), July 1981.

Durkin, D. Comprehension Instruction--Where are You? (No. 1), October 1977. (ERIC Document Reproduction Service No. ED 146 566, 14p., PC-$2.00, MF-$.91)

Durkin, D. What is the Value of the New Interest in Reading Comprehension? (No. 19), November 1980. (ERIC Document Reproduction Service No. ED 198 499, 51p., PC-$5.30, MF-$.91)

Durkin, D. Reading Comprehension Instruction in Five Basal Reader Series (No. 26), July 1981.

Jenkins, J. R., & Pany, D. Teaching Reading Comprehension in the Middle Grades (No. 4), January 1978. (ERIC Document Reproduction Service No. ED 151 756, 36p., PC-$3.65, MF-$.91)

Joag-dev, C., & Steffensen, M. S. Studies of the Bicultural Reader: Implications for Teachers and Librarians (No. 12), January 1980. (ERIC Document Reproduction Service No. ED 181 430, 28p., PC-$3.65, MF-$.91)

McCormick, C., & Mason, J. What Happens to Kindergarten Children's Knowledge about Reading after a Summer Vacation? (No. 21), June 1981.

Osborn, J. The Purposes, Uses, and Contents of Workbooks and Some Guidelines for Teachers and Publishers (No. 27), August 1981.

Pearson, P. D., & Kamil, M. L. Basic Processes and Instructional Practices in Teaching Reading (No. 7), December 1978. (ERIC Document Reproduction Service No. ED 165 118, 29p., PC-$3.65, MF-$.91)

Rubin, A. Making Stories, Making Sense (includes a response by T. Raphael and J. LaZansky) (No. 14), January 1980. (ERIC Document Reproduction Service No. ED 181 432, 42p., PC-$3.65, MF-$.91)

Schallert, D. L., & Kleiman, G. M. Some Reasons Why Teachers are Easier to Understand than Textbooks (No. 9), June 1979. (ERIC Document Reproduction Service No. ED 172 189, 17p., PC-$2.00, MF-$.91)

Steinberg, C., & Bruce, B. Higher-Level Features in Children's Stories: Rhetorical Structure and Conflict (No. 18), October 1980. (ERIC Document Reproduction Service No. ED 198 474, 27p., PC-$3.65, MF-$.91)

Taylor, M., & Ortony, A. Figurative Devices in Black Language: Some Socio-Psycholinguistic Observations (No. 20), May 1981.

Tierney, R. J., & LaZansky, J. The Rights and Responsibilities of Readers and Writers: A Contractual Agreement (includes responses by R. N. Kantor and B. B. Armbruster) (No. 15), January 1980. (ERIC Document Reproduction Service No. ED 181 447, 32p., PC-$3.65, MF-$.91)

Tierney, R. J., Mosenthal, J., & Kantor, R. N. Some Classroom Applications of Text Analysis: Toward Improving Text Selection and Use (No. 17), August 1980. (ERIC Document Reproduction Service No. ED 192 251, 43p., PC-$3.65, MF-$.91)